CLAIMS

1. A method of translating a first schema of data having one structure or semantics into a second schema of data having a second structure or semantics by using an ontology deconstruction and reconstruction transfer mechanism which creates an 'interlanguage' document type definition (interlanguage DTD) in which the interlanguage DTD manages the structure and semantics of the structure and semantics of data to allow an interlanguage definition of the first schema and translation into the second schema by the transfer mechanism.

2. A method according to claim 1 in which the deconstruction and reconstruction transfer mechanism includes: machine-reading tags; interpreting the data format which has been marked up by these tags and detecting its inherent structures or semantics to be included in the interlanguage DTD; and using the interlanguage DTD to transfer the data of the first schema into the second schema.

3. A method according to claim 2 in which the transfer mechanism uses at least two overarching mechanisms.

4. A method according to claim 3 in which the overarching mechanism includes a superordination mechanism and a composition mechanism in which the superordination mechanism constructs tag-to-tag 'is a ...' relationships and the composition mechanism constructs tag-to-tag 'has a ...' relationships.

5. A method according to claim 4 in which within the superordination mechanism, there are any one or more of the submechanisms of hyponymy ('includes in its class ...'), hyperonymy ('is a class of ...'), co-hyperonomy ('is the same as ...'), antonymy ('is the converse of ...') and series ('is related by gradable opposition to ...').

38

6. A method according to claim 4 in which within the composition mechanism, there are any one or more of the submechanisms of meronymy ('is a part of ...'), co-meronymy ('is integrally related to but exclusive of ...'), consistency ('is made of ...'), collectivity ('consists of ...').

7. A method according to claim 4 in which data is imported, and in which supplementary tag-by-tag or field-by-field relationships are generated using one or more of the filter mechanisms of:

taxonomic distance (determining whether the relationships of composition and superordination are too distant to be necessarily valid),

levels of delicacy (determining whether an aggregated data element needs to be disaggregated and re-tagged),

potential semantic incursion (determining identifiable sites of ambiguity), and

translation of silent into active tags or vice versa (determining the level in the hierarchy of composition or superordination at which data needs to be entered to effect superordinate transformations).

8. A method according to claim 7 in which a thesaurus and dictionary is created in combination with the definition of the taxonomy in which the dictionary unpacks the meaning by means of paraphrase and exemplars and the thesaurus display wordings through which meanings can be aptly expressed.

9. A method according to claim 8 in which the method includes mapping a layer of the thesaurus into the paradigm-constituting taxonomy for each schema in a related technology to thereby provide a basis of transfer of data to any other mapped schema in the related technology.

10. A method according to claim 9 in which the thesaurus takes each tagging schema as its starting point, lists its tags and reproduces the definitions and examples as given by each defined tagging schema and against each tag, a direct synonym is provided, whose semantics are coextensive with, or

narrower than, the tag against which the mapping occurs to provide a single equivalent for each mapped tag.

11. A method according to claim 8 in which the dictionary is constructed using five semantic rules: minimised ambiguity; functional clarity; lowest common denominator semantics; the distinction of silent from active tag-concepts; and comprehensive internal cross-reference.

12. A method of providing an apparatus which interpellates source data into a interlanguage DTD format for use in transferring data marked up in a first schema of data having one structure or semantics into a second schema of data having a second structure or semantics including the following steps:

a. providing a quantum of source data of said first schema to a processing and storing apparatus;

b. machine-reading the said source data into an DTD according to a schematic structure of a particular source ontology;

c. automatically reading the structure and semantics ontology immanent in the source data by interpreting this both from the DTD and the way the DTD is realised in that particular instance;

d. applying one or more of the four filters: a delicacy filter, a synonomy filter, a contiguity filter and a subset filter;

e. machine-reading determining from the DTD and its particular instantiation an inherent taxonomic or schematic structure forming the interlanguage DTD comprising of relationships of tags that are unambiguous based on the readable structure of the DTD and evidence drawn from its instantiation in the source data.

13. A method according to claim 12 also including the steps of:

f. providing a structured query for assessment of ambiguous relationships of tags and receiving an assessed response to the structured query to add to the interlanguage DTD.

14. A method according to claim 13 also including the steps of :

g.   drawing implications from the assessed response to the structured query to become part of the memory of the apparatus, for aiding in automatically reading the structure and semantics ontology immanent in the source data.

15. A method according to claim 14 including the following steps:

h.   using the determined interlanguage DTD to build a destination dataset using the mechanisms of superordination (hyponymy, hyperonymy, co-hyperonomy, antonymy and series) and composition (meronymy, co-meronymy, consistency, collectivity).

16. The method of claim 12 wherein part of the process of building a destination data is automated according to the inherent structures readable into the destination ontology or from any previous received assessments of ambiguous structures.

17. A system for translating a first schema of data having one structure or semantics into a second schema of data having a second structure or semantics by using a computer-implemented ontology deconstruction and reconstruction transfer mechanism which automatically creates an 'interlanguage' document type definition (interlanguage DTD) in which the interlanguage DTD automatically manages the structure and semantics of the structure and semantics of data to allow an automatic interlanguage definition of the first schema and translation into the second schema by the transfer mechanism; the system having:

a tag reader for machine-reading tags automatically;

an interpreter for interpreting of data format which has been marked up by the tags and detecting its inherent structures or semantics;

a translator for automatically constructing an interlanguage DTD from the interpreted data format and detected structures or semantics; and

a transfer means for automatically transferring the data of first schema into the second schema via automatic definition of the first schema into the second schema using the interlanguage DTD.

18. A system according to claim 17 in which the interpreter uses the transfer mechanism having at least two overarching mechanisms.

19. A system according to claim 18 in which the overarching mechanism includes a superordination mechanism and a composition mechanism in which the superordination mechanism automatically constructs tag-to-tag 'is a ...' relationships and the composition mechanism automatically constructs tag-to-tag 'has a ...' relationships.

20. A system according to claim 19 in which within the superordination mechanism, there are any one or more of the submechanisms of hyponymy ('includes in its class ...'), hyperonymy ('is a class of ...'), co-hyperonomy ('is the same as ...'), antonymy ('is the converse of ...') and series ('is related by gradable opposition to ...').

21. A system according to claim 19 in which within the composition mechanism, there are any one or more of the submechanisms of meronymy ('is a part of ...'), co-meronymy ('is integrally related to but exclusive of ...'), consistency ('is made of ...'), collectivity ('consists of ...').

22. A system according to claim 19 in which supplementary tag-by-tag or field-by-field queries are automatically generated according to any one or more of the filter mechanisms of:

taxonomic distance (automatically machine-reading whether the relationships of composition and superordination are too distant to be necessarily valid),

levels of delicacy (machine-reading whether an aggregated data element needs to be disaggregated and re-tagged),

potential semantic incursion (machine-reading identifiable sites of ambiguity), and

translation of silent into active tags or vice versa (machine-reading the level in the hierarchy of composition or superordination at which data needs to be entered to effect superordinate transformations).

23. A computer readable medium having a program encoded thereon for translating a first schema of data having one structure or semantics into a second schema of data having a second structure or semantics by using a computer-implemented ontology deconstruction and reconstruction transfer mechanism which automatically creates an 'interlanguage' document type definition (interlanguage DTD) in which the interlanguage DTD automatically manages the structure and semantics of the structure and semantics of data to allow an automatic interlanguage definition of the first schema and translation into the second schema by the transfer mechanism.

24. A computer readable medium according to claim 23 in which a thesaurus and dictionary is provided in combination with the definition of the taxonomy in which the dictionary unpacks the meaning by means of paraphrase and exemplars and the thesaurus display wordings through which meanings can be aptly expressed.

25. A computer readable medium according to claim 24 including a mapping of the thesaurus into the paradigm-constituting taxonomy for each schema in a related technology to thereby provide a basis of transfer of data to any other mapped schema in the related technology.

26. A computer readable medium according to claim 25 in which the thesaurus has each tagging schema as its starting point, lists its tags and provides the definitions and examples as given by each defined tagging schema and against each tag, a direct synonym is provided, whose semantics are coextensive with, or narrower than, the tag against which the mapping occurs to provide a single equivalent for each mapped tag.

43

27. A computer readable medium according to claim 24 in which the dictionary has been constructed using five semantic rules: minimised ambiguity; functional clarity; lowest common denominator semantics; the distinction of silent from active tag-concepts; and comprehensive internal cross-reference.

28. A method for extending the range of useability of ontology driven systems and for creating interoperability between different mark-up schemas for the creation, location and formatting of digital content, the method including the steps of:

a) having a database or datafile of digital content in a Document Type Definition of the first digital mark-up or computer software ontology able to be outputted in a selected format allowed by the first digital mark-up or computer software ontology;

b) organising digital mark-up or computer software tags of the first digital mark-up or computer software ontology into an overarching interlanguage ontology capable of absorbing and incorporating at least one other digital mark-up or computer software ontology;

c) automatically translating a Document Type Definition of the first digital mark-up or computer software ontology into a translated interlanguage Document Type Definition;

d) selecting one of the at least one other digital mark-up or computer software ontology;

e) automatically translating the translated interlanguage Document Type Definition into a Document Type Definition of the selected other digital mark-up or computer software ontology thereby allowing information in the database or datafile format to be output in the required selected format allowed by the selected other digital mark-up or computer software ontology.

29. A method according to claim 28 in which the step of organising digital mark up or computer software tags of the first digital mark-up or computer software ontology into an overarching interlanguage ontology capable of absorbing and incorporating at least one other digital mark-up or computer software ontology includes the steps of indexing according to the following rules:

(i)   providing a first level of granularity such that tags which represent data at a finer level of delicacy in Ontology X produce automatically recomposed data in Ontology Y which manages the same data at a higher level of semantic aggregation.

(ii)   providing a lowest common denominator semantics such that, when data has been data marked up with a pair of tags that can be interpreted to be closely synonymous but not identical, the narrower semantics of the two tags is operationalized.

(iii)   providing contiguous domains wherein tags can be aggregated and aligned by virtue of the fact that they relate to semantically exclusive data.

(iv)   providing subset schemas within a tag such that a whole new domain identified by within Ontology Q or within a defined area of ontology Q can be mapped within a single tag in Ontology R.

30. A method according to claim 29, further comprising returning a set of results from data based on heterogeneous schemas that have been mapped against the interlanguage Document Type Definition.

45